Proton Equilibria in the Minor Groove of DNA

Sue Hanlon,* Linda Wong,# and George R. Pack#
*Department of Biochemistry, University of Illinois College of Medicine, Chicago, Illinois 60612, and #Department of Biomedical Sciences, University of Illinois College of Medicine at Rockford, Rockford, Illinois 61107 USA

ABSTRACT Poisson-Boltzmann calculations by Pack and co-workers suggest the presence of regions of increased hydrogen ion density in the grooves of DNA. As an experimental test of this prediction, we have attached proton-sensitive probes, with variable linker lengths, to random-sequence DNA at G sites in the minor groove. The amino groups of β-alanine, γ-aminobutyric acid (GABA), and ε-aminocaproic acid have been coupled at pH 5, via a formaldehyde link, to the exocyclic amino group of guanine, utilizing a reaction that has been extensively investigated by Hanlon and co-workers. The resulting adducts at pH 5 retained duplex B form but exhibited typical circular dichroism (CD) changes previously shown to be correlated with the presence of a net positive charge in the minor groove. Increases in the solvent pH reversed the CD spectral changes in a manner suggesting deprotonation of the carboxylic acid group of the adduct. These data were used to calculate an apparent pK_a for the COOH. The pK_a was increased by 2.4 units for β-alanine, by 1.7 units for GABA, and by 1.5 units for ε-amino caproic acid, relative to their values in the free amino acid. This agrees well with Poisson-Boltzmann calculations and the energy minimization of the structures of the adducts that place the carboxyl groups in acidic domains whose hydrogen ion density is ~2 orders of magnitude greater than that of bulk solvent.

INTRODUCTION

The relationship between the polyanionic charge distribution of DNA and the molecules and ions of its local environment has profound structural and functional consequences; the theoretical and experimental effects of this relationship have been studied by numerous groups, among them those of both Hanlon and Pack. Theoretical and computational descriptions of counterion condensation have used cylindrical models for DNA with the Poisson-Boltzmann (PB) equation (Schellman and Stigter, 1977; Wilson et al., 1980; Guéron and Weisbuch, 1980; LeBret and Zimm, 1984a), Metropolis Monte Carlo (MC) (LeBret and Zimm, 1984b; Nordenskiold et al., 1984; Pack et al., 1989; Mills et al., 1985), integral equation (Murthy et al., 1985), and Brownian dynamics (Guldbrand, 1989) techniques. These have shown that the mean field approach of the PB equation yields results that are in accord with the more rigorous approaches. The application of the PB equation to a three-dimensional all-atom model with neglect of the dielectric boundary (Klein and Pack, 1983; Pack et al., 1986a,b, 1990) showed that ions cluster in the grooves of duplex DNA. The extension of the PB approach to include a dielectric boundary at the DNA-environment interface corroborated these results (Jayaram et al., 1989; Pack et al., 1993). Comparison of PB and MC with counterion condensation theory (Lamm et al., 1994) showed that these more detailed models of the layer of condensed counterions are consistent with the earlier analysis of Manning (1978).

Calculation of the spatial distribution of hydronium ions using the PB equation led to the prediction of acidic domains within the grooves of DNA (Lamm and Pack, 1990). These calculations revealed that the local density of protons within the grooves can be two orders of magnitude higher than that in bulk solvent. A direct consequence of this model is a position-dependent effect on the pK_a of ionizable groups in the grooves. The negative electrostatic potential arising mainly from the phosphate groups of the backbone substantially increases the value of the activity coefficients of negatively charged groups at the DNA surface. Calculation of the pK_a leads to a higher value (an apparent pK_a) for the group compared to its value in bulk solvent. Because the strength of this field varies within the groove, the activity coefficients vary with the group's position in the groove, and its apparent pK_a is correspondingly affected. According to the acidic domains hypothesis, at physiological ionic strength the apparent pK_a of the carboxyl group of an amino acid bound to the surface of DNA is predicted to increase by approximately the same factor as the local density of counterions—about two pK_a units, depending on the precise location of the titratable site.

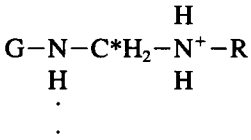
In the present paper we describe an experimental test of this prediction, undertaken in parallel with detailed theoretical calculations for the specific system employed. We have attached a proton-sensitive probe, with variable linker lengths, to random-sequence DNA at guanine bases in the minor groove. Conformational characteristics of the B-form DNA are expected to change as variations in the pH of the bulk solvent alter the charge character of the probe. The relationship between structure and pH can be analyzed, and an apparent pK_a (pK_{APP}) for the attached probe can be determined. This approach, although indirect, is adequate to reveal significant changes in the pK_a of the attached probe.

For these experiments we have utilized an extensive background of information obtained by Hanlon and coworkers on the properties of an adduct of DNA created by

Received for publication 17 July 1996 and in final form 7 October 1996.
Address reprint requests to Dr. George R. Pack, University of Illinois, College of Medicine at Rockford, 1601 Parkview Ave., Rockford, IL 61107. Tel.: 815-395-5694; Fax: 815-395-5666; E-mail: georgeP@uic.edu.

0006-3495/96/01/291/10 $2.00

reacting an aliphatic primary amine and formaldehyde with DNAs of various forms and composition. The results of these studies, reviewed by Maibenco et al. (1989), have revealed that the adduct, in the absence of equilibrium concentrations of amine and formaldehyde, is in a very stable overwound B form in the pH range of 4.5 to 9.0 (Chen et al., 1981, 1983, 1987; Fish et al., 1983; Kilkuskie et al., 1988). Only the G residues are derivatized in the final purified product. The exocyclic amino groups of the G residues are cross-linked to the amino group in a structure shown below:



where the NH·· represents the Watson-Crick hydrogen bond of the 2-amino group of guanine, and the C^*H_2 is derived from formaldehyde.

Although the average change in the winding angle is small (~2°/bp at most, for the most highly derivatized random sequence DNA; Kilkuskie et al., 1988), there is a marked lowering of the rotational strength of the positive band above 260 nm in the circular dichroism (CD) spectrum. This decrease is a linear function of the amount of amine attached to the DNA. It can be demonstrated by reversible titration that the effect on the CD spectrum is due to the presence of the positively charged amino group in the minor groove. Partial removal of the positive charge by titration to pH 10.6 relieves the depression (Chen et al., 1981).

Using this information about the effects of a positive charge at the G locus, we have reasoned that if an amino acid is attached with the COOH protonated, the derivative should exhibit the characteristic depressed positive band above 260 nm seen in the aliphatic amine derivatives. On the other hand, if the amino acid is in the zwitterionic form, there will be no net charge and the positive band of the derivative should have a rotational strength approximating that of the underivatized DNA control. Invoking a two-state model and assuming a linear relationship between the fractional amount of each form and the CD signal, it should be possible to evaluate the pH function, $[COO^-]/[COOH]$, from the changes in the CD signal at fixed wavelength positions in the positive band of the CD spectra. As the pH is increased from values where the COOH is fully protonated to a pH where the COO^- species predominates, the variation in this ratio should equal $\{[\theta]^{pH} - [\theta]^+\}/\{[\theta]^{\pm} - [\theta]^{pH}\}$, where $[\theta]^{\pm}$ and $[\theta]^+$ are the mean residue ellipticities of the zwitterion and the fully protonated forms, respectively, and $[\theta]^{pH}$ is the observed ellipticity of the mixed forms at a given pH. (A similar relationship would obtain for the unnormalized values of the ellipticities, θ, or for any of the other CD functions, ΔE_λ or ΔA). These CD data can thus be used to evaluate the apparent pK_a of each adduct.

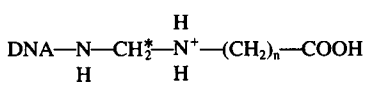
To this end, we have attached several different amino acids, shown in Table 1, to exocyclic NH_2 groups of G residues in random-sequence (calf thymus) DNA. These amino acids differ in the length of the aliphatic carbon chain separating the amino group and the carboxylic acid group. Molecular models suggest that if the amino acid chain were in an extended form, the COOH group of glycine (GLY) would be well within the groove. The COOH group of β-alanine (B ALA) and γ-aminobutyric acid (GABA) would be just within and at the phosphate periphery, respectively. If its carbon chain were extended, the COOH group of ϵ-aminocaproic acid (E CAP) would be well beyond the phosphate perimeter. It was our expectation that as one titrated this series under comparable conditions of solvent and temperature, the apparent pK_as would all be higher than those found in the free amino acids in solution, and would decrease in the order of GLY, B ALA, GABA, and E CAP, relative to the unattached forms. Although the results of our experiments generally conformed to these expectations, there were unanticipated effects that revealed some interesting properties of the minor groove and its interaction with the adduct.

MATERIALS AND METHODS

Preparation and characterization of the amino acid adducts

The preparation of the amino acid adducts followed the general procedure used in previous studies with the *n*-butyl amine adducts (Chen et al., 1981).

TABLE 1 Properties of the amino acid adducts

Stucture of adduct: 

		pK$_a$ COOH at 25°C			
n	Amino acid	AA	Adducts	$\Delta pK_a^{(obsd)}$	$\Delta pK_a^{(calc)}$
1	Glycine (GLY)	2.3	Unstable	—	—
2	β alanine (B ALA)	3.3	5.7 ± 0.2	2.4	1.8
3	γ NH_2 butyric (GABA)	4.3	5.9 ± 0.2	1.7	1.6
5	ϵ NH_2 caproic (E CAP)	4.4	6.0 ± 0.2	1.5	2.0

*(CH_2) is derived from formaldehyde; NH· · · is the portion of the exocylic amino group of guanine involved in a Watson-Crick hydrogen bond with C.
#The pK_as of the COOH groups in the amino acids are perturbed from that of the corresponding carboxylic acids because of the positively charged amino group. The intrinsic pK_a of this group should be ~4.8 at 25°C.

For the labeling experiments, $^{14}CH_2O$ was used to estimate the value of R (= moles amino acid adduct/mole nucleotide) as previously described by Chen et al., (1981) for the amine adducts. The amino acid adducts were prepared in 35 mM NaCl, 10 mM (acetate/acetic acid) at pH 4.8 at room temperature (~25°C). The concentrations of calf thymus DNA, amino acid, and formaldehyde in the reaction mixture were either 0.15 mM (for the titration experiments) or 1.5 mM (for the labeling experiments) in DNA, 10 mM in amino acid, and 0.67 M (2%) in CH_2O. The DNA control contained all components of the reaction mixture except the amino acids. The reactions were followed in either a Jasco Instruments model J600 or a J710 CD spectrometer for up to ~2 h in either 0.1-cm or 1.000-cm path quartz CD cells. Comparable performance of both instruments was routinely checked with standard solutions of d-camphosulfonic acid, as described in the Jasco Instrument manuals. The calf thymus DNA employed in these studies was the high-molecular-weight, highly purified sample used in previous studies (Maibenco et al., 1989; Chen et al., 1981, 1983). The amino acids were the highest purity available from Sigma and Aldrich. All other chemicals and solvents were reagent-grade materials.

As the reaction proceeded, the absorbance spectra of the reaction mixtures were also monitored in the same CD quartz cuvettes over the wavelength range of 220 to 400 nm to detect possible hyperchromic increases and light-scattering artifacts. No significant hyperchromic increases in absorbance were found in these spectra, indicating that little or no denaturation of the duplex samples had occurred in the reaction mixture. Data above 330 nm revealed the absence of light-scattering increases in absorbance as well. Concentrations of DNA in all solutions were estimated from the value of the absorbance at 260 nm using an extinction coefficient of 6700 M^{-1} cm^{-1}, previously determined for this DNA preparation (Chen et al., 1981). These and other absorption spectra were obtained with a Beckman Instruments DU 40 spectrophotometer equipped with Beckman data-capture software.

At the end of the allotted reaction time, the samples were exhaustively dialyzed against the NaCl, acetate buffer, pH 4.8 solvent to remove unreacted amino acids and formaldehyde. The CD spectra of the final dialyzed products were also obtained in one of the two CD instruments at 25°C. Data from both CD instruments were processed with the Jasco data-capture software and the spreadsheet software package Quattro Pro (Borland). All but the GLY adduct were stable to dialysis at pH 4.8 and could be titrated as described below.

In the titration experiments, the pH of the adduct solution was changed by the addition of small aliquots of 1 M NaOH or 1 M HCl to the contents of the CD cell. After obtaining a spectrum, the pH was measured in the cell with a Corning pH meter and micro combination electrode. The possibility of amino acid loss at the elevated pHs was checked by obtaining duplicate spectra 10 min apart for selected pH values. At the alkaline end of the titration, reversibility was evaluated by returning the solutions rapidly to pH 5. As long as the pH remained below 6, there was insignificant loss of the amino acid adduct, and the titration was reversible. At higher pHs, there was a small loss of adduct upon standing at the elevated pH. The data presented here have been analyzed under conditions where the loss of adduct due to elevated pHs is either negligible or has been compensated by omission of the suspect points in the processing of the data.

Calculation of adduct conformations

The first step in the theoretical estimation of the effective pK_a of the amino acid adduct was the determination of an approximate conformation for the system by energy minimization using the AMBER parameter set (Weiner et al., 1986) within HyperChem (Hypercube, 1994) augmented by atomic charges calculated for the amino acids. The free, positively charged *N*-methyl derivatives of the amino acids were geometry-optimized using a semi-empirical molecular orbital method, AM1 (Dewar et al., 1985). The geometry was then held fixed, the carboxylate proton was removed from each of these, and the atomic charges were recalculated for the resulting zwitterionic species. The AM1-derived net atomic charges for the protonated and unprotonated *N*-methyl amino acids were used in the subsequent molecular mechanics and Poisson-Boltzmann calculations.

To generate the structures of the *N*-methyl amino acid–DNA adducts, the B-form of duplex DNA with the sequence d(CGATG*ATCGC)·d(GCGATCATCG) was constructed. The AM1-optimized cationic form of the amino acid was attached to the exocyclic amine of G*, with the loss of one hydrogen atom from the *N*-methyl group and another from the amine. Three classes of orientations of the adduct-DNA system were sampled: the *N*-methyl amino acid aligned toward the 3′ end, toward the 5′ end, and out into solution. Many different orientations around the C-C bonds were also sampled within each of these classes. For each of these starting geometries, the first step was a constrained optimization in which only the *N*-methyl amino acid underwent structural change. For this step, a distance-dependent dielectric constant was used. The lowest energy conformer in each of the three orientational classes was retained. Nineteen Na^+ counterions were generated at the bisectors of the OPO^- angles, to neutralize the system. A periodic box, 6 Å larger in each Cartesian direction, was then constructed around the system and filled with TIP3P water molecules. The positions of the G-C nucleotide pairs at either end of the system were constrained to their starting positions to ensure that the duplex character of the system was retained. All of the remaining atoms in the system were then energy-minimized. A comparison of the energetics of the resulting structures was then done by single-point energy calculations on the *N*-methyl amino acid adduct alone, again using a distance-dependent dielectric constant. The lowest-energy conformation for each of the *N*-methyl amino acids was used for the Poisson-Boltzmann calculations. These conformations are shown in Fig. 1, *A*, *B*, and *C*.

Calculation of the apparent pK_a of the bound probe

Continuum-electrostatic calculations have been employed to calculate pK_a shifts in titratable groups in proteins (Bashford and Karplus, 1990; Bashford and Gerwert, 1992; Yang et al., 1993; Oberoi and Allewell, 1993; Antosiewicz et al., 1994). Similar approaches have been applied to the calculation of salt-dependent pK_a shifts of titratable groups on ligands intercalated into the DNA (Misra and Honig, 1995) using the electrostatic free energy calculated with the formalism derived by Sharp and Honig (1990). Preliminary attempts, using our Poisson-Boltzmann solver, to apply the formalism of Bashford and Karplus (1990) to the current problem showed a strong dependence of predicted pK_a shift on the internal dielectric constant of the macroion. Rather than systematically searching for the best approach to the calculation of pK_a shifts within our methods, we have adopted a less rigorous approach that is less dependent on the specific values assumed for the macroion dielectric constant.

Calculations of the effective pK_a of the carboxylate groups involved application of the nonlinear Poisson-Boltzmann equation to a curvilinear lattice with contours at the adduct-DNA surface, as previously described (Pack et al., 1993). This lattice occupied a cylindrical cell with an axis coincident with the helical axis of the macroion and extended 100 Å radially, yielding a nucleotide concentration of 0.01 M. For each base pair three planar cross sections, perpendicular to the helical axis, with a uniform interplanar separation of 1.13 Å, defined the matrix for the lattice. A rolling-sphere algorithm determined the distance of closest approach for spheres of varying radii. Ions were excluded from the 1.4 Å layer at the innermost surface in the present calculations. This innermost layer and the rest of the environment were assigned a dielectric constant of 78.4; the interior volume defined by the van der Waals surface of DNA was assigned a dielectric constant of 4.

The atomic charges of the adduct-DNA system were mapped to this grid using an algorithm that assigns charge based on the overlap of the atomic van der Waals sphere with the finite volume element of the grid. The distribution of electrolyte charge in the environment was determined by application of the normalized Boltzmann equation:

$$\rho_i^k = \frac{N_k \exp(-\beta z_k \phi_i)}{\Sigma_i v_i \exp(-\beta z_k \phi_i)}, \quad (1)$$

in which ρ_i^k is the density of ion type k in grid element i; N_k is the density of ion type k in bulk solution where the electrostatic potential, ϕ, vanishes;

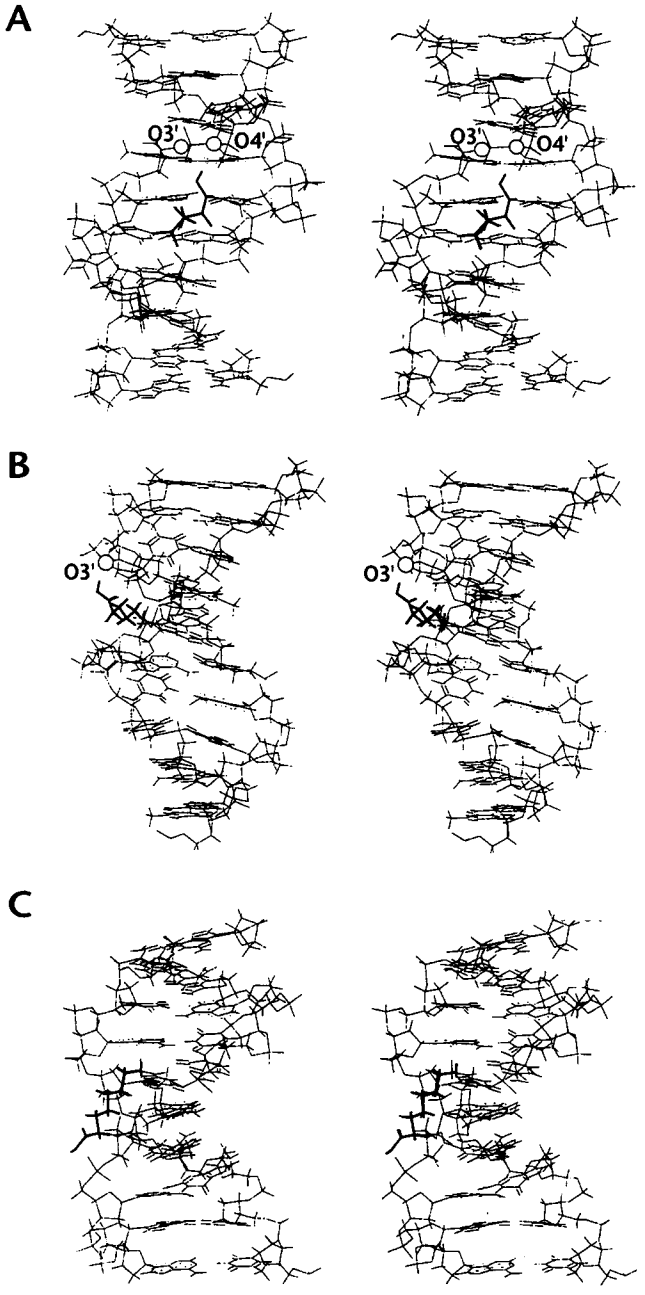


FIGURE 1 Stereo pairs of the optimized conformation of amino acid adducts. Adducts are shown in bold, with the putative H-bonding atoms depicted as circles. (*A*) B ALA; (*B*) GABA; (*C*) E CAP

z_k is the valence of ion type k; and v_i is the volume of grid cell i. The electrostatic potential in grid cell i, ϕ_i, is given by rearranging the finite difference analog of the Poisson equation:

$$\phi_i = \frac{[4\pi v_i \rho_i/\epsilon_i + \Sigma_j(\phi_j \epsilon_{ij} S_{ij}/r_{ij})]}{\Sigma_j(\epsilon_{ij} S_{ij}/r_{ij})}, \qquad (2)$$

in which ϵ_i is the dielectric constant within cell i; S_{ij} is the surface area of the boundary shared by cells i and j; and ϵ_{ij} is the arithmetic average of ϵ_i and ϵ_j. Solution of these coupled equations yielded the electrostatic potential for each finite element within the adduct-DNA boundary and in the external environment.

The activity coefficient for an ion depends on the electrostatic potential; near a polyelectrolyte such as DNA the spatially varying electrostatic potential results in activity coefficients that can be mapped throughout space using electrostatic potentials calculated in the continuum Poisson-Boltzmann approximation (Pack and Wong, 1996; Lamm et al., 1996). The potentials external to the macroion are relatively insensitive to the choice of dielectric constant of the DNA and so can be used to provide an approximate computational scheme for the estimation of the proton equilibria at the DNA surface.

The set of equations relating the experimental data to the value of the apparent dissociation constant of the carboxylate group, K_{APP}, is given below. The ionization in bulk solution,

$$HA \overset{K_a}{\rightleftharpoons} H^+ + A^- \qquad (3)$$

has an acid dissociation constant, K_a, given by

$$K_a = \frac{a_{H+}a_{A-}}{a_{HA}} = \frac{a_{H+}\gamma_-[A^-]}{\gamma_{HA}[HA]}. \qquad (4)$$

When the ionization occurs in the groove region of DNA, K_{APP}, written in terms of the measurable quantities, pH, and the ratio of charged to uncharged acid, becomes

$$K_{APP} = \frac{a_{H+}[A^-]}{[HA]} = \frac{10^{-pH}[A^-]}{[HA]}. \qquad (5)$$

The large magnitude of the negative electrostatic potentials in the grooves lowers the activity coefficients of these groove protons. Because the system is in equilibrium, however, the proton activity (a_{H+}) is spatially invariant; its value in the groove regions can be determined by measuring the bulk pH. The lower activity coefficient and constant activity require that the proton density in the "acidic domains" of the groove regions be increased. To relate the experimentally measurable quantity, K_{APP}, to quantities that can be estimated from continuum electrostatic calculations, we combine Eqs. 4 and 5:

$$K_{APP} = \frac{K_a\gamma_{HA}}{\gamma_-} \approx \frac{K_a}{\gamma_-}. \qquad (6)$$

In Eq. 6 the activity coefficients, γ_{HA} and γ_-, can be approximated by the electrostatic potential in the groove: $\gamma_{HA} = \exp(z_{HA}\beta\phi)$ and $\gamma_- = \exp(z_{A-}\beta\phi)$, with z_{HA} and z_{A-} representing the charges of the protonated and unprotonated carboxylic acid (zero and negative one, respectively), ϕ is the electrostatic potential at the titration site, and $\beta = 1/kT$. The quantity γ_- is a function of the electrostatic potential and hence of position. Similarly, the pK_a shift of a group varies with its position and can be written

$$\Delta pK_a = pK_{APP} - pK_a = -\beta\phi(2.303)^{-1}. \qquad (7)$$

Sharp and Honig (1990) have shown that the free energy change in the nonlinear Poisson-Boltzmann equation is not rigorously a linear function of the electrostatic potential. On the other hand, the simplicity of this approximation and the relative independence of the external electrostatic potential on the parameters chosen to define the macroion and its surface lead us to use this approach to estimate the pK_a shifts.

Analysis of experimental data

Experimental determination of K_{APP} requires an expression for the ratio $[A^-]/[HA]$. As described in the Introduction, the CD signals of the positive band serve as an indicator of the average net charge on the adducts. We

make the assumption that this ratio can be recast as

$$\frac{[A^-]}{[HA]} = \frac{(Y_\lambda - B_\lambda)}{(T_\lambda - Y_\lambda)}. \quad (8)$$

In this equation, Y_λ is the signal of the adduct at a given wavelength, λ, and pH; T_λ is the extrapolated value of the signal at the alkaline extreme; and B_λ is the extrapolated signal at the acid extreme. The units of these signals may either be in ellipticities, θ, in degrees, or in mean residue ellipticities, $[\theta]_\lambda$, in degree cm^2/decimol. To relate these values to the fractions of acid and conjugate base present at a given pH, we assume that the change in the signal is a linear function of the fractional charge borne by the COOH population and that the values of T_λ and B_λ correspond to the signals exhibited by the zwitterion and the fully protonated forms, respectively.

When Eq. 8 is substituted into Eq. 5 and the terms are rearranged, we can define a parameter J that will be a linear function of the CD signal Y. This rearrangement yields

$$J_\lambda = 10^{-pH}(Y_\lambda - B_\lambda) = K_{APP}T_\lambda - K_{APP}Y_\lambda. \quad (9)$$

A plot of the parameter J_λ versus Y_λ will yield the value of K_{APP} and the upper limit, T_λ. We chose this form of the equation, rather than one involving the upper limit T_λ, because B_λ can be more readily evaluated by extrapolation than can the upper limit T_λ. This plot has the disadvantage of mixing two variables and thus making estimation of errors more difficult. It is more sensitive, however, in detecting deviations from a simple two-state model than is the more conventional Scatchard plot, which requires a knowledge of both T_λ and B_λ.

K_{APP} was also determined, however, by attempting to evaluate the upper limit, T_λ, for cases where the change in Y at the elevated pH was not excessive. These calculations involved a least mean-squares determination of the slope and intercept of plots of pH versus $\log[(Y_\lambda - B_\lambda)/(T_\lambda - Y_\lambda)]$. The intercept corresponded to the best fitting value of pK_{APP}. Significant deviations of the slope from the value of 1.0 were taken as indications of cooperativity of proton release.

RESULTS

Computational results

We have calculated the minimum energy structures of the three adducts whose pK_{APP} values could be experimentally determined. The optimized structure of the β-alanine adduct, displayed in Fig. 1 *A*, shows the transferred proton to be buried and directed away from the solvent, suggesting hydrogen bonding between the protonated carboxylate and the DNA. The carboxyl oxygen comes within 3.0 Å of the O3′ (labeled O3*) of the phosphate group in the dC(3′-5′)dG linkage of the opposite strand, and the transferred proton is 2.9 Å from this O3′. A more likely candidate for a hydrogen bond involves the sugar oxygen (labeled O4*) of that dC residue; the O-O distance is 3.1 Å, whereas the H-O distance is 2.2 Å. Efforts to reoptimize the position of the transferred proton always led to this minimum-energy structure.

Fig. 1 *B* shows the lowest energy structure located for the γ-aminobutyric acid adduct. In this structure, the linear chain follows the minor groove, with the carboxylate having direct access to the electrolyte environment; the transferred proton is directed toward the O3′ (labeled as O3*) of a dT residue of the same strand. The carboxylate oxygen-O3′ distance is 2.8 Å, and the H-O3′ distance is 1.9 Å, suggesting hydrogen bonding interactions. In the optimized conformer computed for the ε-aminocaproic acid adduct, shown in Fig. 1 *C*, the methylene chain remained close to the floor of the minor groove, with the protonated carboxylate directed toward the solvent. The chain was not extended, but rather appeared to be embedded in the minor groove. No hydrogen bonding interactions are in evidence here. For all of the calculated conformations there is a slight (~0.2 A) narrowing of the minor groove relative to B-DNA, presumably because of van der Waals interactions of the walls of the groove with the adduct, coupled with the coordinated effects of winding angle increases.

We have calculated ΔpK_a as a function of position for the three minimum-energy conformations shown in Fig. 1, *A*, *B*, and *C*, assuming a concentration of 30 mM added 1-1 monovalent salt present in addition to the 10 mM monovalent cation concentration that neutralized the DNA phosphate charge (i.e., the neutral system contained 40 mM cations and 30 mM anions). Plots of ΔpK_a for the plane in which the transferred proton was calculated to reside are presented in Fig. 2, *A*, *B*, and *C*. The environmental cell closest to the transferred proton is indicated by a white circle; the ΔpKa value in this cell is given in column 6 of Table 1. Although the results presented assume a dielectric constant of 4 for the macromolecule, ΔpK_a calculated in this way is not very sensitive to variations in the dielectric constant. For example, parallel calculations performed with the assumption that DNA had a dielectric constant of 20 lowered this quantity by only 0.05 in each of the three cases presented.

The most striking feature of these maps is that, regardless of conformation, ΔpK_a at the surface was calculated to be in the range 1.4–2.0. Although conformational changes would change the map, the range provides a qualitative measure of the pK_a shifts that can be expected due to the electrostatic potential at the nucleic acid surface. The shifts predicted in the minor groove are generally greater than for the major groove or for the extra-groove regions, in accord with previous descriptions of the acidic domains at the DNA surface (Lamm and Pack, 1990).

These results are in qualitative agreement with the calculations of Misra and Honig, which predict a ΔpK_a of about 2.6 for an intercalated ligand in 30 mM monovalent salt.

Experimental results

Although the apparent rate constants were lower, the reaction of calf thymus DNA with the four amino acids used in this study gave circular dichroism spectra in the reaction solvent that were very similar to those obtained under comparable conditions with the primary amines (Chen et al., 1981, 1983, 1987; Fish et al., 1983; Kilkuskie et al., 1988; Maibenco et al., 1989). All four of the amino acids exhibited pattern changes as a function of time that indicated that product was being formed. The relative rates of reaction were GLY ≪ B ALA < E CAP < GABA. Except for

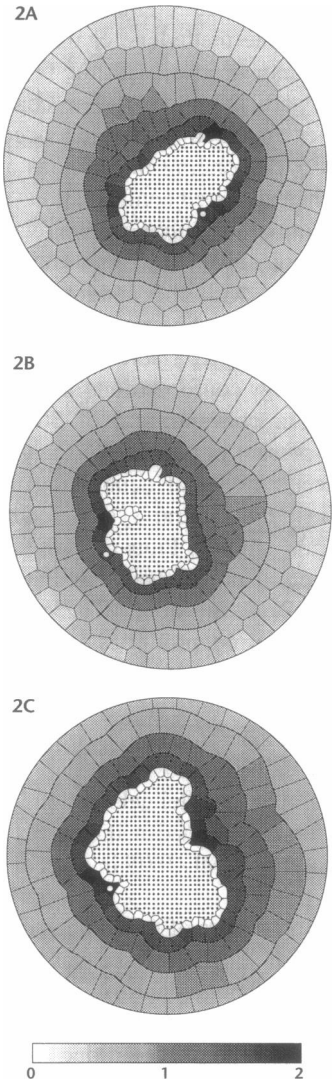


FIGURE 2 Maps of ΔpK_as derived from Eq. 7 and Poisson-Boltzmann electrostatic potentials calculated for the optimized conformation of the amino acid adducts shown in Fig. 1. Dielectric constants of 4 and 78.4 were used for DNA and the environment, respectively; the calculations assumed 30 mM added 1–1 monovalent salt. (*A*) B ALA; (*B*) GABA; (*C*) E CAP

GLY, the CD spectra of the dialyzed products were also similar to those obtained with the primary amines. A set of these dialyzed samples is displayed in Fig. 3 for the DNA control and the four amino acids for a reaction time of ~60 min. The lowering of the rotational strength of the positive band is an effect previously seen in the primary amine spectra, reflecting the magnitude of the amine loading (i.e., moles amine/mole nucleotide) retained upon dialysis. The spectrum of the dialyzed product for the GLY reaction was almost identical to the DNA control, revealing that the reaction product observed with GLY during the course of the reaction did not survive the dialysis procedure. The DNA control solution gave a spectrum that was identical to the unreacted DNA sample, indicating no irreversible denaturation during the reaction.

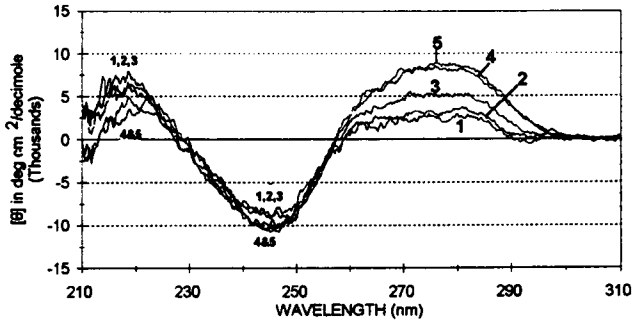


FIGURE 3 CD spectra of the dialyzed amino acid adducts and the underivatized calf thymus DNA control at pH 4.8. (1) GABA; (2) E CAP; (3) B ALA; (4) GLY; (5) the DNA control.

Each of the three products, B ALA, GABA, and E CAP, was then titrated from pH 4.7 to pH ~8-9. Although the spectra changed slowly upon standing at the alkaline extreme, the lower limit on the acid side was quite stable with time. Furthermore, the positive band above 260 nm showed minimal change in the pH range of 4 to 5, suggesting that the adduct was fully protonated at the lower end of the acid range. CD spectral changes in the positive band above 260 nm are shown in Fig. 4, *A*, *B*, and *C*, as a function of pH for B ALA, GABA, and E CAP. Except for spectrum 7, the numbers increase as the pH increases from 4 to ~9, as given in the legend.

The last spectrum, 7, in each panel of these figures is that of the DNA control. It is interesting that at the alkaline end point, where one would expect the carboxyl group to be completely deprotonated, the spectra of the adducts are not identical to that observed in the DNA control. We attribute this to the effect of the quaternary amine linkage immediately adjacent to the bases. Thus there is still a residual positive charge near the guanine that influences the winding angle and, consequently, the CD signal.

In previous studies with the primary amines, the spectral properties of the adduct were linear functions of the extent of derivatization (Chen et al., 1981, 1983; Maibenco et al., 1989). If we were to use the change in $[\theta]_{275}$ to measure the charge difference in the minor groove, it was imperative to demonstrate that this was also true of the adducts created with the amino acids. Adducts of GABA were correspondingly prepared at variable amino acid loadings by removing aliquots of the reaction mixture containing labeled $^{14}CH_2O$ at different times. After dialysis, the samples were counted and the amount of label was translated into a ratio, *R*, expressing moles of amino acid per mole of nucleotide in the product. The spectral properties of this reaction at pH 4.8 were those previously seen in the amine adducts. That is, the rotational strength of the positive band above 260 nm became increasingly depressed with increased amino acid loading, without an appreciable change in the bands at 245 nm and 220 nm. A convenient measure of the rotational strength of this band is the value of $[\theta]$ at 275 nm, approximately the midpoint of this band. Fig. 5 shows the rela-

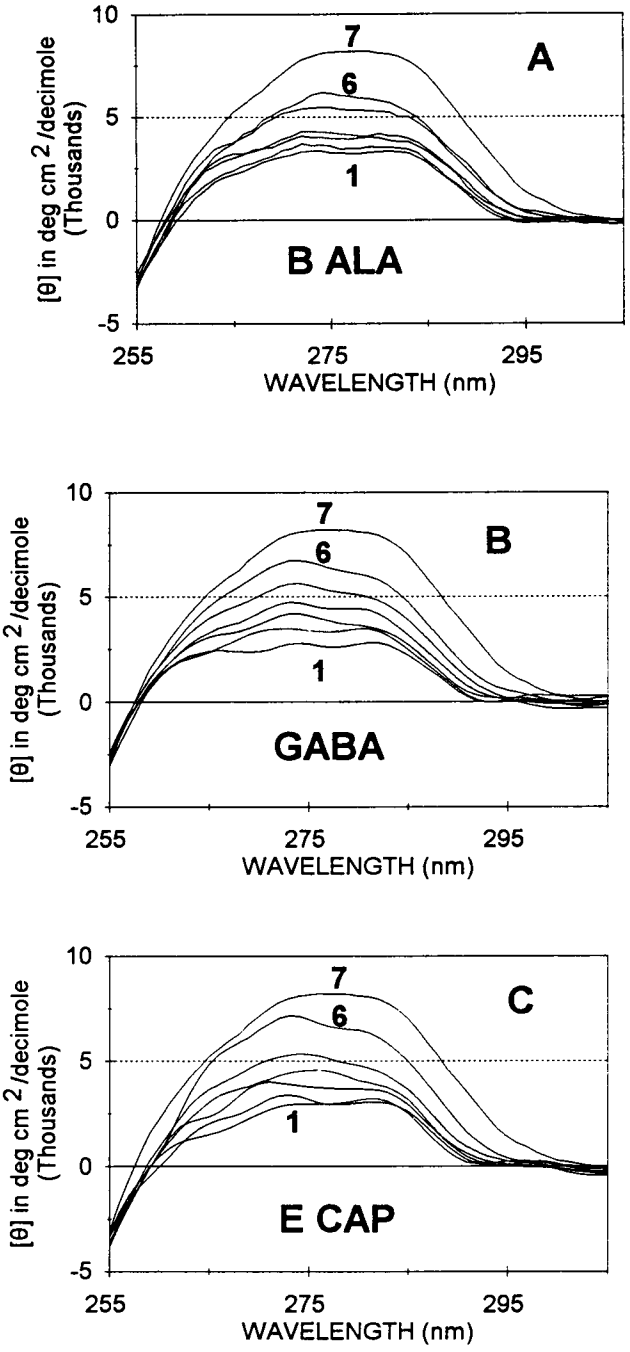


FIGURE 4 Changes in the positive band above 260 nm as a function of pH in the CD spectra of the adducts. Spectrum (7) in each panel is the spectrum of the DNA control at pH 4.8. (*A*) B ALA. pH values are (1) 4.8, (2) 5.1, (3) 5.5, (4) 5.8, (5) 7.1, and (6) 9.0. (*B*) GABA. pH values are (1) 4.65, (2) 5.01, (3) 5.52, (4) 6.02, (5) 6.43, and (6) 9.4. (*C*) E CAP. pH values are (1) 4.3, (2) 4.9, (3) 5.5, (4) 6.0, (5) 6.6, and (6) 9.0.

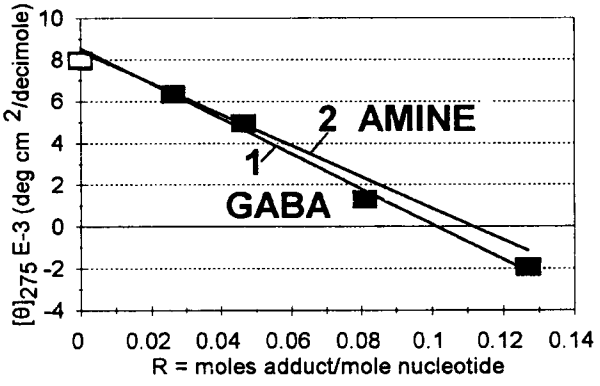


FIGURE 5 Relationship between $[\theta]_{275}$ and the level of substitution of GABA. Line 1 represents the regression of the experimental points (■) for the GABA adducts. Line 2 represents the composite plot for the primary amine adducts examined in previous studies. The point (□) at $R = 0$ is the value for unreacted DNA, not included in either regression analysis.

tionship between the value of $[\theta]_{275}$ and R for both the amino acid adducts and the composite curve for the various primary amines examined in previous studies. The slopes and intercepts are very similar, thus confirming that a linear relationship exists for these amino acid adducts in the present study as well. The close agreement in the values of $[\theta]_{275}$ of the two adducts also confirms our hypothesis that the amino acids are essentially fully protonated at the acidic pH values.

Using the increases in the value of $[\theta]_{275}$ as a measure of the extent of proton removal from the adduct, the data for β-alanine, γ-NH_2-butyric acid, and ε-NH_2 caproic acid were analyzed, as described in Materials and Methods, in two ways. First, the values of $[\theta]_{275}$ at the acid and alkaline ends of the titration range were used as end points to calculate the log ratio of (A/HA) for each titration experiment. The results of the regression analysis of the pH versus log (A/HA) plot were then transformed into the display of $[\theta]_{275}$ versus pH. These plots, showing both the results of the regression analysis (*solid lines*) and the observed data (*points*) are displayed in Fig. 6, *A*, *B*, and *C*, for B ALA, GABA, and E CAP. The two curves for B ALA in Fig. 6 *A* represent the behavior of the same solution in the initial forward titration (curve 1) and a retitration (curve 2) after standing for 10 min at pH 9. The pK_{APP} for each titration is the same (5.7), but the curves are displaced by an increment along the ordinate that presumably reflects the loss of B ALA adduct at the elevated pH. In all three cases, the regression analyses indicated no cooperative proton loss.

The other analysis procedure used Eq. 9 and made no assumptions about the stability of the upper limit, T_λ. These plots of the function J versus $[\theta]_{275}$ are shown in Fig. 7, *A*, *B*, and *C*, for B ALA, GABA, and E CAP, respectively. Although the data were scattered, there was no evidence of nonlinearity. The pK_{APP}s evaluated from the slopes did not differ significantly from the more conventional analysis described above. Omission of the last point shown in each panel from the regression resulted in a change of slope amounting to no more than a difference of 0.1 in pK_{APP}. The pK_{APP} values determined by the regression analyses of all points are given in the fourth column of Table 1.

The analyses described above were performed with adducts whose loading of amino acids was low (≤0.06 mole amine/nucleotide). At the higher loading levels for GABA

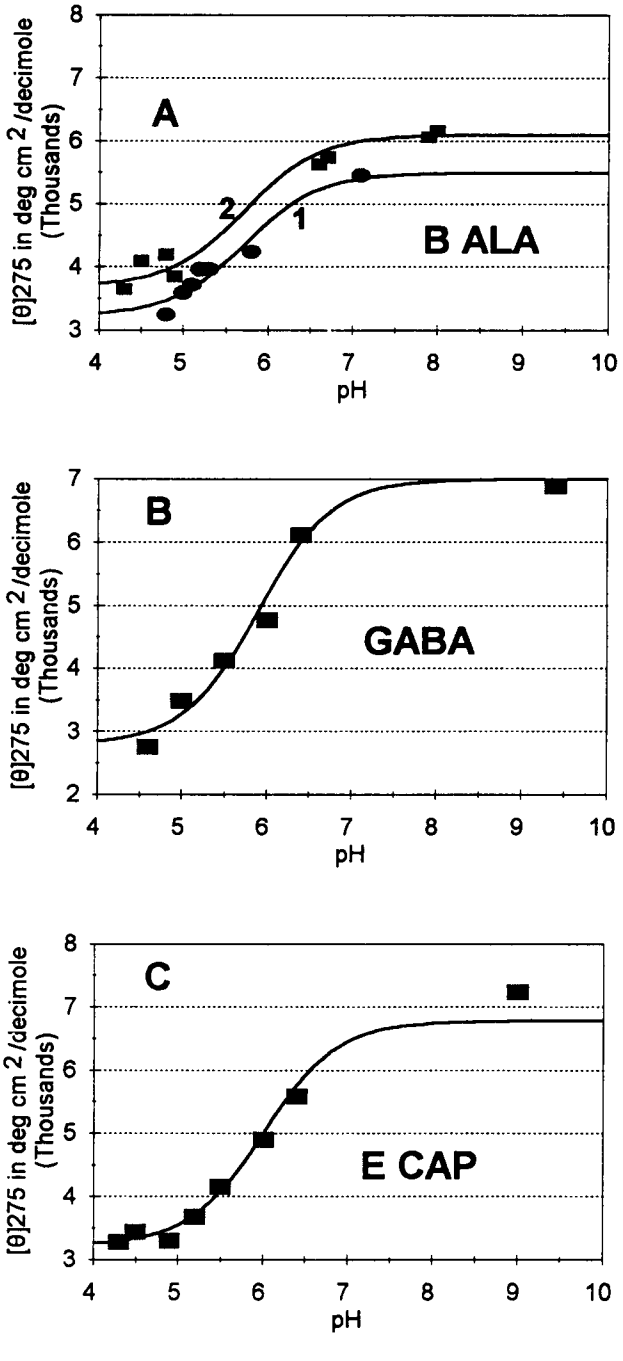


FIGURE 6 Changes in the values of $[\theta]_{275}$ as a function of pH for the three adducts, B ALA (*A*), GABA (*B*), and E CAP (*C*). The solid line in each panel represents the hypothetical curve generated as described in the text. In *A* (B ALA), curve 1 represents the original solution. Curve 2 represents a retitration of the solution after standing at pH 9 for 10 min before returning to pH 4.3.

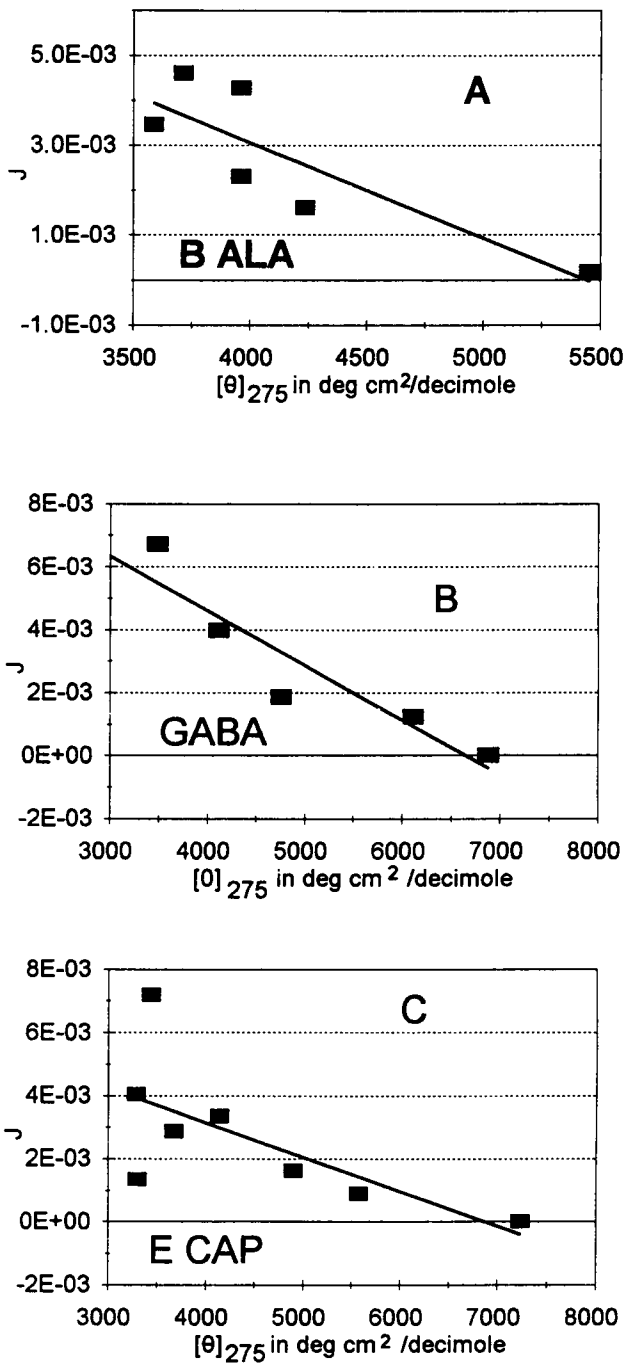


FIGURE 7 Plots of the *J* function, defined in Eq. 9, against $[\theta]_{275}$. (*A*) B ALA, (*B*) GABA, and (*C*) E CAP.

and E CAP (0.10 to 0.12), however, the *J* plots were distinctly nonlinear, and the regression slopes in the plots of pH versus log(A/HA) were significantly greater than 1. The data for GABA indicated that the pK_{APP} decreased monotonically from 6.6 to 5.9 as protons were removed by titration. The last value of 5.9 corresponds to the average value for this adduct at the lower level of loading.

This pK_{APP} behavior is the reverse of what might be anticipated from a secondary electrostatic effect caused by removing protons from the carboxyl group. This effect may have several origins. In part, it may result from the uncertainty of values of T_λ (the basic end point) due to exposure to the elevated pH. Although we cannot rule this out entirely, this uncertainty is unlikely to contribute significantly, because the extrapolated pK_{APP} for the last proton to dissociate is essentially that evaluated for the adducts at the

lower level of substitution. The more likely explanation is that there is a conformational or solvation effect caused by overloading of the adduct that results in greater stability of the protonated form of the amino acid. Once proton removal begins, the structure formed at the fully protonated state is destabilized, and this effect may be transmitted to near neighbors, resulting in the facilitated loss of protons from the adjacent sites. It should be noted, however, that this conformational effect does not affect the linear relationship between $[\theta]_{275}$ and R, as our data in Fig. 5 extend to this range of ellipticity values for which this pK_{APP} behavior exists.

Because of the observed nonlinearity, an uncertainty has been introduced into the evaluation of the pK_{APP}s of the more highly derivatized adducts. We have thus restricted our analysis to the adducts of lesser loading that show no appreciable effects of this type. The data shown in column 4 of Table 1 are for the lesser substituted adducts ($R \leq 0.06$). The fifth column in this table gives the observed values of ΔpK_a (= $pK_{APP} - pK_a$ of the free amino acid). The last column shows the theoretical values of ΔpK_a calculated by the procedures described in the Calculation section.

Considering the assumptions upon which these analyses rest, the agreement is quite good. The magnitude of the deviations is not excessive and can be readily accounted for. Because of the longer chain length, it was our initial expectation that the ϵ-NH_2 caproic acid adduct would extend farther from the macroion surface and so exhibit a significantly lower ΔpK_a than that of GABA. As seen in Table 1, this was not so. An examination of the minimum energy structures for the adducts in their fully protonated states, shown in Fig. 1, *A*, *B*, and *C*, indicates that the methylene chain of the caproic acid derivative was not extended into the solution, but rather was wrapped around the minor groove. The position of its COOH group is comparable to that in the GABA adduct and thus resides in a locus at or within the periphery of the PO_4 groups of the DNA. The β-alanine moiety is fairly well buried in the groove and the proton is involved in a hydrogen bond, as noted above. This is expected to raise the apparent pK_a, as discussed in a previous section.

CONCLUSIONS

The apparent pK_as of the COOH groups on the amino acid adducts are substantially higher than those for the same amino acids free in solution. Although these results are qualitatively similar to those predicted from the simple acidic domains model of Lamm and Pack (1990), there are other contributions to the pK_a shifts. Hydrogen bonding of the carboxyl groups to the phosphate oxygens is seen in the calculated structure for the β-alanine adduct. The stabilization energy that results from the electronic redistribution within the hydrogen bond is considered in neither the acidic domains model nor in other continuum-based electrostatic descriptions of pK_a shifts. That stabilization energy may be responsible for the relatively large shift of 2.4, compared to the calculated value of 1.8. Changes in the macromolecular dynamics and concomitant dielectric constant effects that the simple model does not consider may affect the results in ways discussed by others (Antosiewisz et al., 1994). Although the dielectric constant of the macroion only minimally changes our predictions of pK_a shifts, motions of the attached group take it through regions of varying activity coeffients. In the case of the ϵ-NH_2 caproic acid adduct, it can be seen from Fig. 6 that a small excursion of the carboxylate would take it to a region where ΔpK_a would change from 2 to 1.8 or 1.6. It seems likely that such conformations contribute to the observed value of 1.6.

The combination of experiment and theory presented here allows us to unequivocally state that the protonic equilibria in the minor groove, at least, are significantly different from that in bulk solvent. As a result, acidic and some basic groups that normally would be unprotonated under physiological conditions may be protonated. This is important for our understanding of the mechanisms of the reactions and interactions in which DNA participates in vivo.

We thank Drs. A. S. Benight and T. Keiderling for the use of the CD facilities at the Department of Chemistry, UIC. We are also indebted to the Research Resource Center, UIC, for the use of their CD instrumentation.

This work was supported in part by grant GM29070 to GRP from the National Institutes of Health and funds from the Campus Research Board at the University of Illinois at Chicago.

REFERENCES

Antosiewicz, J., J. A. McCammon, and M. K. Gilson. 1994. Prediction of pH-dependent properties of proteins. *J. Mol. Biol.* 238:415–436.

Bashford, D., and K. Gerwert. 1992. Electrostatic calculations of the pKa values of ionizable groups in bacteriorhodopsin. *J. Mol. Biol.* 224: 473–486.

Bashford, D., and M. Karplus. 1990. pKa's of ionizable groups in proteins: atomic detail from a continuum electrostatic model. *Biochemistry.* 29: 10219–10225.

Chen, C., R. Kilkuskie, and S. Hanlon. 1981. Circular dichroism spectral properties of covalent complexes of deoxyribonucleic acid and *n*-butylamine. *Biochemistry.* 20:4987–4995.

Chen, C. Y., B. H. Pheiffer, S. B. Zimmerman, and S. Hanlon. 1983. Conformational characteristics of deoxyribonucleic acid-butylamine complexes with C-type circular dichroism spectra. 1. An X-ray fiber diffraction study. *Biochemistry.* 22:4746–4751.

Chen, C., S. Ringquist, and S. Hanlon. 1987. Covalent attachment of an alkylamine prevents the B to Z transition in poly(dG-dC). *Biochemistry.* 26:8213–8221.

Dewar, M. J. S., E. G. Zoebisch, E. F. Healy, and J. J. P. Stewart. 1985. AM1: a new general purpose quantum mechanical molecular model. *J. Am. Chem. Soc.* 107:3902–3909.

Fish, S. R., C. Y. Chen, G. J. Thomas Jr., and S. Hanlon. 1983. Conformational characteristics of deoxyribonucleic acid-butylamine complexes with C-type circular dichroism spectra. 2. A Raman spectroscopic study. *Biochemistry.* 22:4751–4756.

Guéron, M., and G. Weisbuch. 1980. Polyelectrolyte theory. I. Counterion accumulation, site-binding, and their insensitivity to polyelectrolyte shape in solutions containing finite salt concentrations. *Biopolymers.* 19:353–382.

Guldbrand, L. 1989. The distribution and dynamics of small ions in simulations of ordered polyelectrolyte solutions. *Mol. Physiol.* 67: 217–237.

Hypercube, Inc. 1994. HyperChem for Windows: Release 4. Waterloo, Ontario, Canada.

Jayaram, B., K. A. Sharp, and B. Honig. 1989. The electrostatic potential of B-DNA. *Biopolymers.* 28:975–993.

Kilkuskie, R., N. Wood, S. Ringquist, R. Shinn, and S. Hanlon. 1988. Effects of charge modification on the helical period of duplex DNA. *Biochemistry.* 27:4377–4386.

Klein, B. J., and G. R. Pack. 1983. Calculations of the spatial distribution of charge density in the environment of DNA. *Biopolymers.* 22: 2331–2352.

Lamm, G., and G. R. Pack. 1990. Acidic domains around nucleic acids. *Proc. Natl. Acad. Sci. USA.* 87:9033–9036.

Lamm, G., L. Wong, and G. R. Pack. 1994. Monte Carlo and Poisson-Boltzmann calculations of the fraction of counterions bound to DNA. *Biopolymers.* 34:227–237.

Lamm, G., L. Wong, and G. R. Pack. 1996. DNA-mediated acid catalysis: calculations of the rates of DNA-catalyzed hydrolyses of diol epoxides. *J. Am. Chem. Soc.* 118:3325–3331.

LeBret, M., and B. H. Zimm. 1984a. Distribution of counterions around a cylindrical polyelectrolyte and Manning's condensation theory. *Biopolymers.* 23:287–312.

LeBret, M., and B. H. Zimm. 1984b. Monte Carlo determination of the distribution of ions about a cylindrical polyelectrolyte. *Biopolymers.* 23:271–285.

Maibenco, D., P. Tang, R. Shinn, and S. Hanlon. 1989. Base and conformational specificity of an amine modification of DNA. *Biopolymers.* 28:549–571.

Manning, G. S. 1978. The molecular theory of polyelectrolyte solutions with applications to the electrostatic properties of polynucleotides. *Q. Rev. Biophys.* 2:179–246.

Mills, P., C. F. Anderson, and M. T. Record, Jr. 1985. Monte Carlo studies of counterion-DNA interactions. Comparison of the radial distribution of counterions with predictions of other polyelectrolyte theories. *J. Chem. Phys.* 89:3984–3994.

Misra, V. K., and B. Honig. 1995. On the magnitude of the electrostatic contribution to ligand-DNA interactions. *Proc. Natl. Acad. Sci. USA.* 92:4691–4695.

Murthy, C. S., R. J. Bacquet, and P. J. Rossky. 1985. Ionic distributions near polyelectrolytes. A comparison of theoretical approaches. *J. Phys. Chem.* 89:701–710.

Nordenskiöld, L., D. K. Chang, C. F. Anderson, and M. T. Record, Jr. 1984. ^{23}Na NMR relaxation study of the effects of conformation and base composition on the interactions of counterions with double-helical DNA. *Biochemistry.* 23:4309–4317.

Oberoi, H., and N. M. Allewell. 1993. Multigrid solution of the nonlinear Poisson-Boltzmann equation and calculation of titration curves. *Biophys. J.* 65:48–55.

Pack, G. R., G. A. Garrett, L. Wong, and G. Lamm. 1993. The effect of a variable dielectric coefficient and finite ion size on Poisson-Boltzmann calculations of DNA-electrolyte systems. *Biophys. J.* 65:1363–1370.

Pack, G. R., and B. J. Klein. 1984. Generalized Poisson-Boltzmann calculation of the distribution of electrolyte ions around the B- and Z-conformers of DNA. *Biopolymers.* 23:2801–2823.

Pack, G. R., G. Lamm, L. Wong, and D. Clifton. 1990. The structure of the electrolyte environment of DNA. *In* Theoretical Biochemistry and Molecular Biophysics, Vol. 1. D. L. Beveridge and R. Lavery, editors. Adenine Press, New York. 237–246.

Pack, G. R., C. V. Prasad, J. S. Salafsky, and L. Wong. 1986a. Calculations on the effect of methylation on the electrostatic stability of the B- and Z-conformers of DNA. *Biopolymers.* 25:1697–1715.

Pack, G. R., and L. Wong. 1996. Electrostatic effects on the rates of DNA-catalyzed reactions. *Chem. Phys.* 204:279–288.

Pack, G. R., L. Wong, and G. Lamm. 1989. Model systems for DNA and its environment: suitability and accuracy in theoretical calculations. *Int. J. Quant. Chem.* 16:1–15.

Pack, G. R., L. Wong, and C. V. Prasad. 1986b. Counterion distribution around DNA: variation with conformation and sequence. *Nucleic Acids Res.* 14:1479–1493, 5547.

Schellman, J. A., and D. Stigter. 1977. Electrical double layer, zeta potential, and electrophoretic charge of double-stranded DNA. *Biopolymers.* 16:1415–1434.

Sharp, K. A., and B. Honig. 1990. Calculating total electrostatic energies with the nonlinear Poisson-Boltzmann equation. *J. Phys. Chem.* 94: 7684–7692.

Weiner, S. J., P. A. Kollman, D. T. Nguyen, and D. A. Case. 1986. An all-atom force field for simulations of proteins and nucleic acids. *J. Comp. Chem.* 7:230–252.

Wilson, R. W., D. C. Rau, and V. A. Bloomfield. 1980. Comparison of polyelectrolyte theories of the binding of cations to DNA. *Biophys. J.* 30:317–325.

Yang, A.-S., M. R. Gunner, R. Sampogna, K. Sharp, and B. Honig. 1993. On the calculation of pKas in proteins. *Proteins Struct. Funct. Genet.* 15:252–265.